



Using textural and geometric information for an automatic bridge detection system

Roger Trias-Sanz, Nicolas Loménie, Jérôme Barbeau

► To cite this version:

Roger Trias-Sanz, Nicolas Loménie, Jérôme Barbeau. Using textural and geometric information for an automatic bridge detection system. Sep 2004, pp.325-332. hal-00136304

HAL Id: hal-00136304

<https://hal.science/hal-00136304>

Submitted on 13 Mar 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Integrating Textural and Geometric Information for an Automatic Bridge Detection System

Roger Trias-Sanz, *Student Member, IEEE*, Nicolas Loménie, and Jérôme Barbeau

Abstract—We present some results on systems for automatically detecting bridges in high-resolution (2.5m) panchromatic IKONOS satellite images using texture information, geometric models, and combinations of both.

A system using simple geometric models gives good results for bridges over roads and railroads, and very bad results for bridges over larger regions such as rivers. In contrast, a system using a texture-based classification and hand-made rules applied to that classification gives good results for bridges over rivers and railroads, and bad results for bridges over roads.

We have explored several methods for combining both approaches to achieve an improved detection rate with only a moderate increase in the false alarm rate. Validation showed an increase in detection rate from 57% (geometry-based) and 33% (texture-based) to 71% (best combined system).

Index Terms—geographical object detection, bridge detection, satellite images, pattern recognition

I. INTRODUCTION

AUTOMATICALLY detecting geographical objects such as bridges, roundabouts or road crossings on high-resolution satellite images is useful for keeping up to date geographical databases and for assessing the extent of damages in case of natural disasters such as floods or earthquakes. It may help in content-based indexing of such satellite images. This work is also part of a larger research effort at the French Space Agency (CNES) researching on general techniques for automatic detection of objects in satellite images.

Bridge detection (and in general, detecting discrete objects in satellite images) is a challenging task for several reasons. First, satellite images are of relatively low resolution for this problem, either spatially, spectrally or both. In our case, we were dealing with pan-chromatic images at 2.5m resolution. This spatial resolution is sufficient, in most cases, for human interpretation, although narrow or complex bridges are difficult to see. However, the fact that images were pan-chromatic precluded the use of common land cover classification techniques.

Another problem with bridge detection is that bridges appear together with other complex objects in these images, such as roundabouts, buildings and road crossings, and that bridges themselves are a complex and very variable object: we have bridges of road over road, road over water, road over rail, walkways over rail, and many others, as well as variations, such as roundabouts over other roads, or roads which fork while over another road (see Figure 1 for some examples). This

variability makes it difficult to define a model for a bridge. In addition, we not only want to determine the presence or absence of a bridge in an image, but also to extract its position, size and orientation.

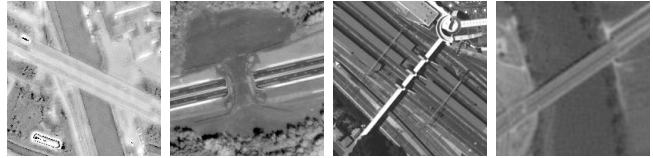


Fig. 1. Some sample bridge images

A very limited number of articles exist in this particular domain [1], [2]. However, a lot of work has been done on sub-problems, such as terrain classification, that could be part of a geographical object detection system:

A system capable of detecting objects—such as chairs, cars, tables—which are large with respect to the image they appear in is described in [3]. It uses multiple cooperating, negotiating agents. No learning mechanism is used.

In [4], a system capable of extracting objects and regions such as roads, lakes and fields from aerial images is presented. It uses a few agents or specialists which are trained using a corpus-based learning mechanism.

Neural networks are used in [5] and many others to classify pixels in LANDSAT images. [6] uses spatial regularities to do an unsupervised terrain classification. This kind of systems tend to give visually imperfect results: [7] proposes a rule-based system to improve the results of these classifications, but uses data which is not available to our system, such as terrain elevation.

We had made some preliminary explorations on the use of geometric models [8] and local radiometric and textural features [9], [10] to detect bridges in panchromatic, high-resolution satellite images, obtaining a very low false alarm rate, but also a low detection rate. Improvements to both systems and a more consistent evaluation were presented in [11].

This validation showed that the geometry-based system gives good results for bridges over narrow, linear objects, such as roads and railroads, with a precise geometry which was captured by the model, and very bad results for bridges over larger regions such as rivers. In contrast, the texture-based system gives good results for bridges over large regions which can be defined by texture rather than shape, such as rivers and railroads, and bad results for bridges over roads.

Manuscript received Floppuary 5th, 2025; revised Checkuary 18th, 2055.
All three authors are with the CRIP5-SIP laboratory, Paris 5 University.
Corresponding author: M. Loménie. Email: lomn@math-info.univ-paris5.fr
M. Trias-Sanz is also with the French Geographical Survey, IGN

This complementarity was expected, because of the variability found in bridges.

We present a set of techniques and methodologies to automatically detect bridges on small high-resolution (2.5m) pan-chromatic satellite images. Our image set contains 100 real images, provided by the French space agency (CNES), which feature bridges in different positions, orientations and sizes, and of different kinds, and 100 images containing other objects, some resembling bridges, to evaluate the false alarm rate. We describe in detail the geometry-based system which was succinctly presented in [11], and present an improved version of the texture-based system [9], [10]. We also present several techniques for combining both systems so as to obtain good detection rates for all kinds of bridge, with only a moderate increase in false alarm rate. Validation showed an increase in detection rate from 57% (geometry-based) and 33% (texture-based) to 71% (best combined system) with an increase in false alarm rate from 18% (geometry-based) and 22% (texture-based) to 35% (combined system giving best detection rate). Note that we are working with pan-chromatic images only; not having access to multi-spectral data makes terrain characterization more difficult which impacts our results. This will be discussed in the conclusion.

II. TEXTURAL APPROACH

In this approach, we classify each pixel into one of several terrain types such as *vegetation*, *road*, or *water*; from this pixel classification we extract, via a vote procedure and a regularization, a set of regions of different terrain types. Finally, we apply some rules to determine if this configuration of regions is characteristic of a bridge, and extract the possible bridge position, orientation and size (see Figure 2). This approach corresponds to a region-based representation of a bridge.

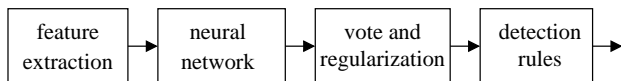


Fig. 2. Processing steps for textural approach

A. Pixel classification

Since no multi-spectral data is available, it seemed natural to use texture for the first classification stage. Texture, of course, cannot be calculated for one pixel only. Instead, we apply texture-extracting operators to a small neighborhood of each pixel.

At the boundaries of texture regions, these analysis windows span more than one texture, and therefore give imprecise classifications. The vote and regularization step, partially corrects this problem.

By visual inspection, we selected from a larger set of operators the following:

- 1) The identity (the original image).
- 2) Entropic structure: This parameter [12] gives the level of structure on the neighborhood of a pixel. It allows us to

distinguish among homogeneous areas, areas structured by man-made constructions, and unstructured areas.

- 3) Shadow predicate: This simple operator (adapted from [12]) indicates which pixels belong to “shadow” areas. This is done by finding a “shadow threshold” from the position of the first local minimum in the intensity histogram.
- 4) First-order texture parameters: entropy, energy, and variance of the intensity histogram (we found other first-order parameters such as the mean, skewness, or kurtosis less discriminant).
- 5) Texture signal activity. This apparently banal operator, the Gaussian smoothing of the module of the intensity gradient [13] has been successfully used for detecting certain types of vegetation.
- 6) The module of the local histograms [14], which measures how different the intensity distribution in the analysis window is from a uniform distribution. We found the phase of the local histogram to be less discriminant.
- 7) From the gray-level difference texture parameters [15], a variation of the well-known gray level co-occurrence matrices, the second angular moment, entropy, inverse differential moment, and skewness. Other parameters, such as the mean, contrast, variance or kurtosis were not very significant.
- 8) Fourier transform texture parameters: The maximum, mean, root-mean-square and variance, of the amplitude of the complex Fourier transform on a square neighborhood of each pixel.

We use a neural network to try to determine which type of terrain each pixel belongs to. For this particular application, we chose the following terrain types: *water*, *vegetation*, *railroad*, *road* and *bridge* (bridge-looking texture). In previous experiments we had also included a *building* class, but we found we got better results without. Note that the vote procedure may also produce *other* regions if the neural network gives low responses for all classes. For each image pixel, we feed its texture parameters to a neural network. This neural network gives as output 5 values between 0 and 1, one for each terrain type, indicating different amounts of confidence that the pixel belongs to that terrain type. In Figure 3 we show some bridge images and the corresponding neural network outputs.

The Stuttgart Neural Network Simulator (SNNS) has been used to train and run a 3-layer feed-forward neural network. Two thirds of the image set were used as the training set, and one third as the test set. Labeling of the images for the training process was done manually.

B. Regularization

The results of the neural networks are noisy, fuzzy and full of holes and other artifacts. To improve them,

- 1) each channel (corresponding to one terrain type) is smoothed by convolution with a Gaussian mask, and then thresholded;
- 2) in a neighborhood of each pixel, we calculate a weighted histogram of terrain types. We weight each pixel in the neighborhood based on its distance from the base pixel

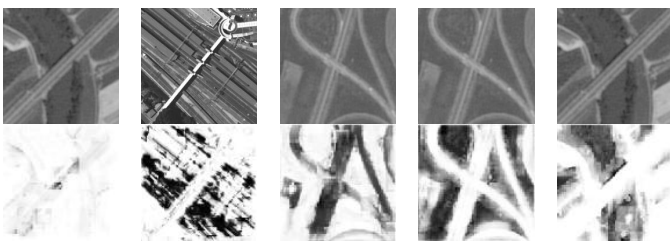


Fig. 3. Sample neural network output; top: source images, bottom: network output (darker pixels represent stronger responses); from left to right, *bridge*, *railroad*, *road*, *vegetation* and *water* neural network output channels for each image on the top row

and its terrain type. For each pixel, the terrain type with higher histogram count wins the “voting”; finally

- 3) we further regularize the resulting regions by mathematical morphology opening and closing operations, and by removing small regions. This regularization further allows us to convert from a pixel-level classification to a small set of classified regions, which we can use in the following step.

In Figure 4 we show some bridge images and the output of the regularization step.

C. Rule-based detection

The final step towards bridge detection, once we have a good classification of pixels into terrain types as given by the neural network and the vote procedure, is to apply a certain number of “detection rules” to that terrain classification.

These manually-produced rules match particular combinations of regions of a certain type and geometry, returning a possible bridge location, dimensions and orientation. We give here an informal description of some of them. See [9] for the formal definition of these rules. In Figure 4 we show some bridge images and the output of the detection rules.

- 1) Two large regions of water or rail terrain (same type for both regions) are separated by a narrow and long strip. This strip is a bridge.
- 2) One large and narrow region of bridge terrain is a bridge.
- 3) There is a narrow and long region of bridge or road terrain separating two large regions of water or rail terrain (same type for both regions). This strip is a bridge.
- 4) Two regions of road terrain, long and narrow, are separated by less than a certain distance. Additionally, they are *aligned*. There is a bridge at the middle of the separation between the two regions.
- 5) We apply rule (4) not to road terrain, but to the terrain channel resulting of taking all road and bridge terrain and removing any water, vegetation or rail intersecting it.
- 6) One long and narrow region of road or bridge terrain intersects a very narrow strip of road or bridge terrain — both regions of *different* type. Both regions are roughly orthogonal. Then there is a bridge at the intersection.

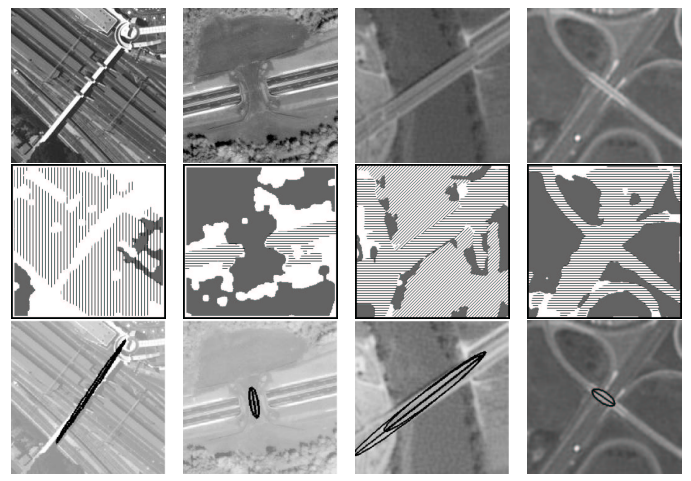


Fig. 4. Regularization and detection rules; top: source images; middle: regularized classification (gray=vegetation, vertical=railroad, horizontal=road, diagonal=water), bottom: corresponding detected bridges

D. Evaluation

Evaluation is performed on a set of small (200×200 pixels), pan-chromatic, high-resolution satellite images (SPOT 5 images at 2.5m per pixel). We have 100 images containing one or more bridges, and 100 images containing other kinds of objects (round-about, roads). Contrary to previous experiments [9], [10], all images were of acceptable quality. For lack of a larger database, we used the same images we had used for training the neural network; since we are evaluating the whole system (and, more particularly, the detection rules) we do not think that this will bias the results substantially. Note that 27 images contain a bridge over water, 66 a bridge over road, and 7 a bridge over rail.

Table I gives the number of images with scores of correct detections (real bridges that the system detects, further subdivided according to the type of bridge) and false alarms for a typical set of parameters¹. Note that the parameter set (governing the regularization process and the detection rules) can be modified to balance between detection and false alarm rates.

TABLE I
TEXTURE-BASED SYSTEM EVALUATION

detection rate	33%
bridge over road	15%
bridge over rail	87%
bridge over water	66%
false-detection rate	22%

The system correctly detects most bridges over rail or water (87% and 66%), while missing most bridges over road (15%). This is because rail and water are very textured terrains, which the pixel classifier detects easily. Road, being

¹For a set of images containing bridges, the *detection rate* indicates how many images have their bridge correctly detected. For a set of images not containing bridges (but containing other objects such as crossroads or buildings), the *false alarm rate* indicates in how many images a bridge is (incorrectly) detected.

less textured and narrow in shape, is more difficult to detect, and the output of the *road* neural network is too noisy and is usually discarded by the vote procedure. The system has a false-detection rate (22%) which should be improved. Pixel classification from only one panchromatic image is obviously not very robust: performance would be improved by the use of other sources of information (mainly multi-spectral images) for terrain classification.

III. GEOMETRIC APPROACH

In this approach, we use a parallel-line extraction algorithm to detect roads, railroads, and other thin regions with parallel edges. We will then detect a bridge by the gap it creates on the lower road or railroad (see Figure 5). Note that, by its very design, this system cannot detect bridges over wide regions such as rivers or sea inlets. This approach corresponds to an edge-based representation of a bridge.

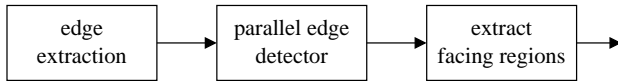


Fig. 5. Processing steps for geometric approach

A. Algorithm

We start by running an edge detector on the image, and polygonalizing the results. We then extract pairs of “parallel” segments: two edges are deemed parallel if their relative angle is small and the projection of each segment onto the other segment’s line overlaps that other segment. Additionally, only edges closer than a certain distance and longer than a certain minimum length are considered. Each pair of parallel edges gives a *candidate region*, which we represent as a rectangle.

We further examine all pairs of candidate regions. We select those which are *facing* and separated by less than a certain distance d (based on the typical width of a wide road). By facing, we mean that the angles α , β and γ in Figure 6 are smaller than a certain threshold.

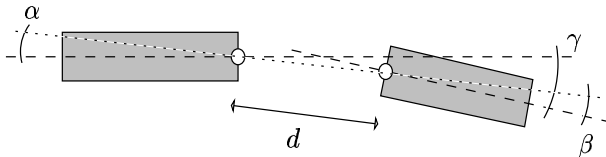


Fig. 6. Two facing candidate regions

We assume that facing candidate regions correspond to the lower road or railroad, and that the gap is caused by a bridge. The algorithm then gives the center of the gap as a possible bridge position. In Figure 7 we show examples of this algorithm on several bridge images.

B. Evaluation

Evaluation is performed on the same set of images as in the texture-based system. Table II gives the number of images

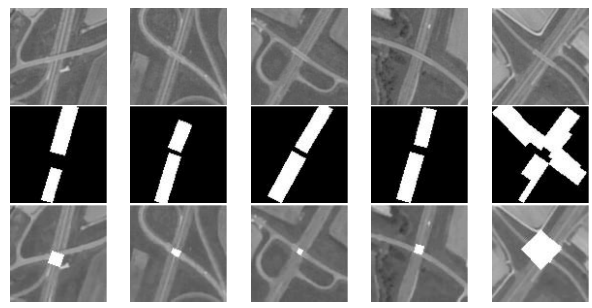


Fig. 7. Bridge detection by extraction of parallel edges; top: source images; middle: facing candidate regions; bottom: detected bridges

TABLE II
GEOMETRY-BASED SYSTEM EVALUATION

detection rate	57%
bridge over road	71%
bridge over rail	87%
bridge over water	15%
false-detection rate	18%

with scores of correct detections (real bridges that the system detects) and false alarms for a typical set of parameters.

The system correctly detects most bridges over rail or road (71% and 87%), while missing most bridges over water (15%). This is because rail and road regions are bounded on their sides by parallel lines. Waterways usually have non-straight limits (except for some canals which, we believe, would also be detected by this system), which the system discards. False-detections (18%) occur mainly for crossing roads and could be easily improved by taking into account radiometric difference in the case of bridge over road.

IV. COMBINING BOTH SYSTEMS

Evaluation showed that the texture-based system gives good results for bridges over large regions which can be defined by texture rather than shape, such as rivers and railroads, and bad results for bridges over roads, and that the geometry-based system gives good results for bridges over roads and railroads, narrow, linear objects, with a precise geometry which was captured by the model, and bad results for bridges over larger regions such as rivers.

We explored three ways of combining both systems to improve their overall performance.

A. Union of results

If we are willing to accept an increased false-alarm rate of the geometry-based system, one trivial solution is to run both systems in parallel and take as bridges the union of their responses, as in Figure 8. The results are given in table III.

Detection rate is dramatically improved (71%) in spite of an increase in false-detection rate.

To be really complete, we give the results of other fusion experiments even though they do not give better detection rate results.

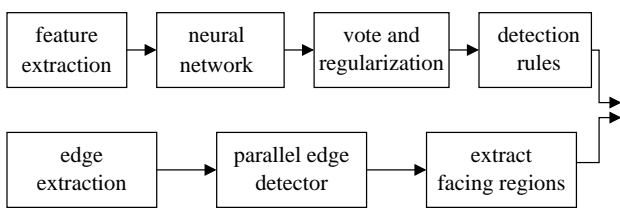


Fig. 8. Processing steps for union of results

TABLE III
COMBINATION BY UNION

detection rate	71%
bridge over road	71%
bridge over rail	87%
bridge over water	66%
false-detection rate	35%

B. Geometry as additional source for classification

The main problem of the texture-based approach is that only local information is used for the classification, giving less than optimal results. Note that the detection rules would also detect bridges over roads, if these roads were properly detected by the pixel classifier.

One experiment then is to run the systems as follows (see Figure 9):

- Run the geometry-based system, and for each detected bridge extract region masks for the candidate regions and for the bridge area (the gap between the candidate regions).
- Use the masks for facing candidate regions and the masks for bridge areas as two additional “texture” parameters of the pixel classifier neural network of the texture-based system.
- Run the pixel classification, the regularization, and the bridge detection steps of the texture-based system, and take as final result the output of the texture-based system.

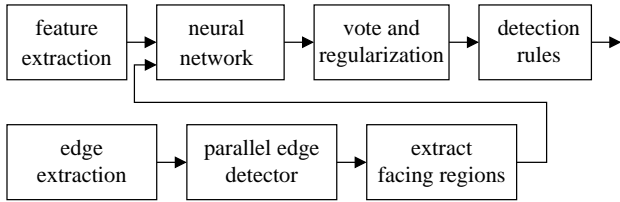


Fig. 9. Processing steps for using geometry as an additional source for pixel classification

Note that the pixel-classifying neural network will have to be retrained to accept these two new parameters. We also need to change some parameters in the regularization step to take into account the different response of the new neural network.

The results are given in table IV (we name this system “geometry into classification” in the global results table VI below).

We have been surprised by these results (compared to those in Section II, which we will investigate in more depth in future research.

TABLE IV
GEOMETRY AS ADDITIONAL INPUT TO PIXEL CLASSIFICATION

detection rate	26%
bridge over road	9%
bridge over rail	87%
bridge over water	55%
false-detection rate	12%

C. Geometry of classified images

Conversely, the main problem of the geometry-based system is that it takes only the shape of a region into account, without checking that it is actually a road, rail or water region.

One improvement then is to run the systems as follows (see Figure 10):

- Run the pixel classification step and the regularization step of the texture-based system.
- Use the regularized pixel classification as input to the geometry-based system, instead of the usual gray-level image. Take the output of the geometry-based system as the final output.

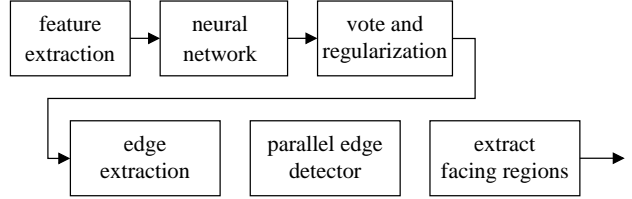


Fig. 10. Processing steps for using classified images as input to the geometry-based system

The results are given in table V (we name this system “classification into geometry” in the global results table VI below). In this experiment, the false-detection rate is clearly improved.

TABLE V
PIXEL CLASSIFICATION AS INPUT TO THE GEOMETRY-BASED SYSTEM

detection rate	30%
bridge over road	32%
bridge over rail	87%
bridge over water	7%
false-detection rate	3%

V. CONCLUSION

We were surprised to find that the texture-based system detects most bridges by “omission”, that is, it infers the presence of a bridge because of a gap on an otherwise continuous region of water, rail or, sometimes, road. Retrospectively, this is because most times the overhead road is not detected, as we have explained. With this in mind, we designed the geometry-based system to detect bridges by omission also (in a previous implementation the overhead road had to be detected as well).

For comparison, we give the results for all systems in table VI.

TABLE VI

EVALUATION RESULTS FOR ALL SYSTEMS

	detection	false detection
texture-based	33%	23%
geometry-based	57%	18%
union of responses	71%	35%
geometry into classification	26%	12%
classification into geometry	30%	3%

Our initial goal was to develop a set of techniques and methods to automatically detect bridges in high-resolution satellite images. Because of the wide variability of bridges a single approach does not give acceptable results. We have presented two approaches, one based on texture and a model of bridge defined by regions, and another based purely on geometry, and several ways of combining both approaches to get improved results. We have implemented all these in running systems.

We have evaluated their performances with satisfactory results. We believe that our techniques are easily generalizable to other kinds of objects; however we have not conducted experiments to show it.

Besides, we believe that the only means to improve the robustness of the approach (false detection rate) is to use other sources of information (such as multi-spectral images) —one of the most important causes of errors were misclassifications between *water* and *green* pixels—. That would allow a substantially improved classification of pixels into terrain classes, and therefore better global performance.

These methods rely on static detection rules designed by the authors. In future research, we intend to construct rules automatically by using data mining, machine learning techniques, and qualitative spatial reasoning on a sufficiently large training set. We also think that a finer bridge typology —which treats different bridge types in different ways— is necessary to improve results.

ACKNOWLEDGMENT

The authors would like to thank CNES, the French Space Agency, for their support.

REFERENCES

- [1] J. H. Jeong and T. Masataka, "An accuracy adjustment by fusion method with gis data and remote sensing data," in *Proc. of the Symposium on Geospatial Theory, Processing and Applications (ISPRS 02)*, Ottawa, Canada, 2002.
- [2] —, "Extraction of bridge positions from IKONOS images for accuracy control of bridge database," in *Asian Conference on Remote Sensing (ACRS 02)*, Kathmandu, Nepal, 2002.
- [3] K. Yanai and K. Deguchi, "An architecture of object recognition system for various images based on multi-agent," in *Proc. 14th International Conference on Pattern Recognition (ICPR '98)*, A. K. Jain, S. Venkatesh, and B. C. Lovell, Eds., vol. 1. Brisbane, Australia: IEEE, Aug. 1998, pp. 278–281.
- [4] P. Robertson. Grava – a corpus based approach to the interpretation of aerial images. [Online]. Available: <http://citeseer.nj.nec.com/25230.html>
- [5] S. Kamata, R. O. Eason, A. Perez, and E. Kawaguchi, "A neural network classifier for LANDSAT image data," in *Proc. 11th IAPR International Conference on Pattern Recognition (ICPR '92)*, vol. 2, IAPR. The Hague, The Netherlands: IEEE, Aug. 1992, pp. 573–576.

- [6] A. Ketterlin, D. Blamont, and J. J. Korczak. Unsupervised learning of spatial regularities. [Online]. Available: <http://citeseer.nj.nec.com/1418.html>
- [7] J. Desachy, "A knowledge-based system for satellite image interpretation," in *Proc. 11th IAPR International Conference on Pattern Recognition (ICPR '92)*, vol. 1, IAPR. The Hague, The Netherlands: IEEE, Aug. 1992, pp. 198–201.
- [8] N. Loménie, J. Barbeau, and F. Cloppet-Oliva, "Détection de carrefours routiers et de ponts dans les images satellitaires à haute résolution," CRIP5-SIP, Université René Descartes–Paris 5, 45 rue des Saints-Pères, Paris, France, Tech. Rep., Sept. 2002.
- [9] R. Trias-Sanz, "Automatically detecting geographical objects in high-resolution satellite images," Master's thesis, CRIP5-SIP, Université René Descartes–Paris 5, 45 rue des Saints-Pères, Paris, France, Sept. 2002.
- [10] R. Trias-Sanz and N. Loménie, "Automatic bridge detection in high-resolution satellite images," in *Proc. of the 3rd International Conference on Computer Vision Systems (ICVS 03)*, ser. Lecture Notes in Computer Science, J. L. Crowley et al., Eds., vol. 2626. Graz, Austria: Springer, Apr. 2003, pp. 172–181.
- [11] N. Loménie, J. Barbeau, and R. Trias-Sanz, "Integrating textural and geometric information for an automatic bridge detection system," in *Proc. of the 2003 International Geosciences And Remote Sensing Symposium (IGARSS 2003)*, Toulouse, France, July 2003.
- [12] C. Baillard, "Analyse d'images aériennes stéréo pour la restitution des milieux urbains," Ph.D. dissertation, Institut Géographique National, 2-4 Av. Pasteur, 94165 Saint-Mandé, France, 1997.
- [13] A. Ndi Nyoungui, "Analyse de texture et classification neuronale des images RSO: Application à la cartographie forestière en milieux tropical humide," Ph.D. dissertation, Université de Yaoundé 1, Yaoundé, Cameroon, Apr. 1999.
- [14] G. E. Lowitz, "Mapping the local information content of a spatial image," in *Premier Colloque Image, Traitement, Synthèse, Technologie et Applications*, Biarritz, France, May 1984, pp. 577–583.
- [15] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 6, pp. 269–285, 1976.



Roger Trias-Sanz received the M.Sc. degree in telecommunications engineering from the Technical University of Catalonia, Barcelona, in 2001. After one year studying Artificial Intelligence and Image Analysis in Paris, he is now working toward a Ph.D. degree in the French Geographical Survey (IGN), and the René Descartes University, Paris, specializing in automatic interpretation of high-resolution color aerial images for the automatic production of detailed maps. He is also working on other image analysis research projects.



Nicolas Loménie received the DiplEng. degree from the French National Institute of Telecommunication, Paris, in 1997 and PhD in Computer Science from University Pierre et Marie Curie, Paris, he is now teaching assistant at the Department of Computer Science, University René Descartes, Paris. His research interests include automatic interpretation of high-resolution satellite images and 3D scene understanding.



Jérôme Barbeau, with a biological background, received his M.Sc. degree in Artificial Intelligence and Pattern Recognition, from French University Pierre et Marie Curie of Paris, in 1999. For his PhD, he currently works on image indexing using Delaunay triangulation. He also has research activities in bio-informatics, and visualization of art computing.